

TD 2 : Prédiction de la météo à J+7 avec des réseaux RNN et CNN

Enzo DAMION et Armelle COUTAUX

October 28, 2022

En réutilisant le code du cours, nous avons implémenté un réseau de neurone récurrent qui prédit la météo 7 jours après une date donnée, en s'appuyant sur la météo actuelle et les prédictions précédentes. La base d'apprentissage est constituée des données journalières de météo pour l'année 2019 et nous testons la validité et précision de notre modèle sur les données de 2020. Il s'agit d'un apprentissage supervisé.

Comparaison du modèle RNN et CNN :

Les deux modèles utilisent le critère de *mean square error* (**MSE**) comme erreur du modèle et le score R2 pour évaluer les performances du modèle.

1. Le modèle **RNN** implémente le `rnn` de la librairie `pytorch` puis un simple perceptron à 10 couches cachée (`mlp`).
2. Le modèle **CNN** réalise les prédictions sans récurrence à l'aide d'une simple convolution en dimension 1 (`cnn` avec un kernel de 3 et un stride de 1) suivi d'un perceptron à une couche cachée de 10 neurones (`mlp`).

Analyse des résultats

En réalisant des test avec des taux d'apprentissage différent pour le RNN (taux d'apprentissage de 0,001 pour la Figure RNN 1 et 0,0001 pour la figure RNN 2), on constate que le modèle converge atteint une erreur faible plus rapidement avec un taux d'apprentissage plus faible. Cependant, le score R2 est relativement faible (environ 0,6). En doublant le nombre de neurones dans la couche cachée du `mlp` final (20 neurones), on obtient une convergence plus rapide et un R2 score légèrement supérieure (Figure RNN 3)

En comparant avec le modèle du CNN, on constate que ce dernier converge moins rapidement que le RNN, bien qu'il montre des bons résultats. En effet, il faut environ trois fois plus d'époques au CNN qu'au RNN pour arriver à la même valeur d'erreur faible : 600 vs 200 époques, comme le montre la Figure CNN 1. Cependant, l'entraînement du réseau CNN est bien plus rapide que celui du RNN en terme de temps : pour 10 000 époques, 21,2 secondes pour le CNN contre 260 secondes pour le RNN.

On observe de l'overfitting à partir de 8000 époques environ pour le RNN, avec un learning rate de 0,0001. L'erreur passe en effet de 0,007543 à 0,007556 entre les époques 9 000 et 10 000, donc une reprise de 140 millièmes (voir Figure RNN overfit). Pour le modèle du CNN, on observe un overfit léger à partir de 3 500 époques.